

Probing the mechanism of the quantum speed-up by time-symmetric quantum mechanics

Giuseppe Castagnoli

Via San Bernardo 9/A, 16031 Pieve Ligure, Italy
giuseppe.castagnoli@gmail.com

June 19, 2012

Abstract

We consider a major family of quantum algorithms. Bob chooses a function out of a set of functions and gives Alice the black box that computes it. Alice is to find a characteristic of the function chosen by Bob (e. g. the period) by performing function evaluation for different values of the argument. The time-symmetric representation of these quantum algorithms shows that: (i) the characteristic of the function found by Alice is a by-product of her reconstruction of Bob's choice and (ii) since there is maximal quantum correlation between Bob's choice and Alice's reconstruction of it, all is as if Alice by reading the reconstruction at the end of the algorithm determined half of Bob's choice. This quantum retroaction of the output on the input explains why quantum algorithms create a stronger than classical input-output correlation, viz. the quantum speed-up. This explanation allows us to pinpoint the exact mechanism of the quantum speed-up. Given a set of functions, this mechanism – under maximization of a quantity of information – becomes the quantum algorithm that yields a characteristic of the function chosen by Bob. This should be a break through, until now there was no known mechanism for producing quantum speed-ups.

1 Foreword

We propose an explanation of the mechanism underlying the quantum speed-up – quantum algorithms requiring fewer computation steps than the corresponding classical algorithms.

We try and position the present proposal within state of the art. It is fair to say that, twenty-seven years after the discovery of the seminal quantum algorithm on the part of Deutsch [1], there is not yet an exact explanation of the mechanism of the quantum speed-up. By this we mean a precise description of the entire mechanism that produces the speed-up. For example, the notion of quantum parallel computation, although has inspired all the quantum

algorithms found so far, cannot be considered the description of a mechanism. It only says that the simultaneous computation of a quantum superposition of inputs and quantum interference are essential ingredients of the speed-up. Although various relations between speed-up and other special quantum features, like entanglement and discord, have been pinpointed in special situations, none of them can be considered a description of the full mechanism of the speed-up.

A related problem still requiring an answer is that of the fundamental nature of the speed up. Apropos of this, it is important to note that some quantum algorithms, like the seminal Deutsch's algorithm and Grover's [2] algorithm, require fewer computation steps than the minimum demonstrably required in the classical case. Now, there is a striking contrast between the amount of literature produced on the quantum violation of limits applying to classical processes like Bell's inequalities and no literature at all on the same kind of violation on the part of quantum computation.

The present explanation of the speed-up relies on the time-symmetric [3] quantum mechanics of Aharonov et al. – see also Vaidman's update of this theory [4] – and on the retro-causal nature of quantum measurement highlighted by partial measurement, as from the work of Dolev and Elitzur [5] on the non-sequential behavior of the wave function.

2 Summary

We consider an important family of quantum algorithms. Bob chooses a function out of a set of functions and gives Alice the black box that computes it. Alice is to find a characteristic of the function chosen by Bob by performing function evaluation for different values of the argument.

Let us focus on a simple example of speed-up. Bob hides a ball in one of four drawers. Alice is to locate it by trying different drawers. If the algorithm is classical, to be sure of locating the ball, Alice should plan to open three drawers. In Grover's quantum algorithm [2], only one drawer suffices, what is obviously the quantum violation of a classical limit.

The time-symmetric representation of these quantum algorithms requires the physical representation of both Alice's action of solving the problem and Bob's action of setting the problem (e. g. by choosing the number of the drawer with the ball). In this representation, the quantum process is considered in its canonical entirety – initial measurement, unitary transformation, final measurement. See Figure 1, where S. stands for state, M. for measurement, U is the unitary part of the quantum algorithm.

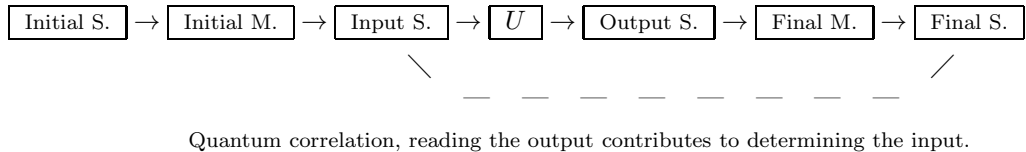


Fig. 1 Time-symmetric representation of quantum algorithms

As we will see, this representation highlights the quantum correlation existing between Bob's choice and the solution found by Alice – see the dotted line in the figure. Because of it, all is as if Alice, by reconstructing Bob's choice, determined half of it. This retroaction of the output on the input is what allows the quantum algorithm to build up a stronger than classical input-output correlation – viz. the quantum speed-up.

The plan of the paper is as follows:

Section 3.1 *Time-symmetric representation*. We develop our explanation on Grover's algorithm. To the usual *Alice's register*, containing the number of the drawer that Alice wants to open, we add a *Bob's register*, containing the number of the drawer with the ball. The initial state of Bob's register is maximally mixed, so that Bob's process of choice is represented from scratch; Bob measures the content of his register obtaining a drawer number at random – Figure 1. We assume that Bob's choice is this very number. The corresponding eigenstate, with the usual sharp state of Alice's register, is the input of the unitary part of the quantum algorithm U . Here, by performing function evaluations, Alice reconstructs Bob's choice in her register. By finally measuring the content of this register, she acquires the number of the drawer chosen by Bob, namely the solution of the problem. There is maximal quantum correlation between the result of the final measurement performed by Alice and that of the initial measurement performed by Bob; in fact the two measurements provide a common, randomly selected, eigenvalue (Bob's choice). We will see that quantum correlation remains there also when Bob unitarily changes the initial random measurement outcome into a desired number.

Section 3.2 *Sharing the determination of Bob's choice*. The fact that measuring an observable determines one of its eigenvalues is a postulate of quantum mechanics. However, this postulate gives room to an ambiguity. When two completely redundant measurements determine a common eigenvalue (here the number of the drawer chosen by Bob), which of the two measurements determines it? Answering this question is instrumental to explaining the quantum speed-up. The idea that the determination should be ascribed to the measurement performed first does not seem to be justified. Because of redundancy, Bob's measurement can be suppressed and the determination of Bob's choice is performed by Alice's measurement, also at the time of the suppressed measurement. In fact, the projection of the quantum state due to Alice's measurement can be advanced at the time of the suppressed measurement along the unitary evolution between the two measurements (by the inverse of the time-forward evolution). Since either measurement determines Bob's choice, for reasons of symmetry we postulate that this determination shares evenly between the two measurements. This means that Alice's measurement contributes to the initial choice of Bob by determining half of it ($n/2$ of the bits that specify it in the case that Bob's choice is an unstructured n -bit string). This explains all the quantum speed-ups examined in this paper.

Section 3.3 *Advanced knowledge*. The time-symmetric representation of the quantum algorithm has another important property. It can be seen as the usual representation, starting with a well determined choice of Bob, relativized to

the observer Alice in the sense of relational quantum mechanics [6]. In this interpretation, the maximally mixed initial state of Bob’s register physically represents Alice’s initial ignorance of Bob’s choice. By advancing the contribution of Alice’s measurement to the determination of Bob’s choice to the time of this choice, the entropy of the initial state of Bob’s register is halved. This means that Alice knows in advance half of the choice.

Section 3.4 *History superposition picture*. Correspondingly, the quantum algorithm is a superposition of histories. In each of them, Alice knows in advance one of the possible halves of Bob’s choice ($n/2$ bits) and performs the classical computations required to identify the missing half. This requires $O(2^{n/2})$ function evaluations against the $O(2^n)$ of the classical case where Alice completely ignores Bob’s choice, what explains the quadratic speed-up of the present algorithm.

Section 4 *Mechanism of the quantum speed-up*. The present explanation of the quantum speed-up allows us to pinpoint the exact mechanism of the quantum speed-up. Given any set of functions, this mechanism – under maximization of a quantity of information – becomes the quantum algorithm that reconstructs Bob’s choice with a quantum speed-up. Yielding a characteristic of the function chosen by Bob is a by-product of this reconstruction.

Section 5 *Deutsch’s algorithm*, Section 6 *Simon’s and the hidden subgroup algorithms*. Here Bob’s choice is a highly structured bit string. In all the cases examined, given the advanced knowledge of half of it, finding the missing half requires a single function evaluation – against an exponential number thereof in the absence of advanced knowledge. This explains the exponential speed-up of these latter algorithms.

Section 7 *Conclusions*. We summarize and discuss the results obtained.

An earlier version of the present work has been presented at the 92nd Annual Meeting of the AAAS Pacific Division [7]. With respect to the explanation of the speed-up provided in Ref. [8], we have brought to the fundamental level the problem of sharing between Alice’s and Bob’s measurements the determination of Bob’s choice. This has allowed us to extend that explanation to the entire family of quantum algorithms.

3 Grover’s algorithm

We develop our argument in detail for the four drawer instance of Grover’s algorithm.

3.1 Time-symmetric representation

We formalize the four drawer problem. Let $\mathbf{b} \equiv b_0b_1$ and $\mathbf{a} \equiv a_0a_1 \in \{00, 01, 10, 11\}$ be the number of the drawer with the ball and respectively that of the drawer that Alice wants to open. Bob writes his choice of the value of \mathbf{b} in a two-qubit register B . As we will see, this register is indifferently real or imaginary. We take the expression “imaginary register” from reference [9], which highlights the

problem-solution symmetry of Grover's and the phase estimation algorithms. Alice writes a value of \mathbf{a} in a two-qubit register A . Then a black box computes the Kronecker function $\delta(\mathbf{b}, \mathbf{a})$, which gives 1 if $\mathbf{b} = \mathbf{a}$ and 0 otherwise – tells Alice whether the ball is in drawer \mathbf{a} . A one-qubit register V is meant to contain the result of the computation of $\delta(\mathbf{b}, \mathbf{a})$ – modulo 2 added to its former content for logical reversibility.

We assume that register B is initially in a maximally mixed state, so that the value of \mathbf{b} is completely undetermined. Further below we will show that this assumption just yields a special view of the usual quantum algorithm (starting with the value of \mathbf{b} completely determined). Registers A and V are prepared as required by Grover's algorithm. In the present time-symmetric representation, the initial state of the three registers is:

$$|\psi\rangle = \frac{1}{2\sqrt{2}} (e^{i\varphi_0} |00\rangle_B + e^{i\varphi_1} |01\rangle_B + e^{i\varphi_2} |10\rangle_B + e^{i\varphi_3} |11\rangle_B) |00\rangle_A (|0\rangle_V - |1\rangle_V). \quad (1)$$

For reasons of space, we use the random phase representation of a density operator – a formalism implicit in the random phase approximation [10]. The φ_i are independent random phases, each with uniform distribution in $[0, 2\pi]$. The density operator is the average over all φ_i of the product of the ket by the bra:

$$\begin{aligned} \langle |\psi\rangle \langle \psi| \rangle_{\forall \varphi_i} = & \frac{1}{8} (|00\rangle_B \langle 00|_B + |01\rangle_B \langle 01|_B + |10\rangle_B \langle 10|_B + |11\rangle_B \langle 11|_B) \\ & |00\rangle_A \langle 00|_A (|0\rangle_V - |1\rangle_V) (\langle 0|_V - \langle 1|_V). \end{aligned} \quad (2)$$

The von Neumann entropy of state (1) is two bits.

Reading the content of a register amounts to measuring a corresponding observable. We call \hat{B} (\hat{A}) the content of register B (A), of eigenvalue \mathbf{b} (\mathbf{a}). \hat{B} and \hat{A} are both diagonal in the computational basis and thus commute. In order to prepare register B in the desired value of \mathbf{b} , in the first place Bob should measure \hat{B} in state (1). He obtains an eigenvalue at random, say $\mathbf{b} = 01$. Correspondingly, state (1) is projected on:

$$P_B |\psi\rangle = \frac{1}{\sqrt{2}} |01\rangle_B |00\rangle_A (|0\rangle_V - |1\rangle_V), \quad (3)$$

we denote projection operators by the letter P . For the time being, we assume that Bob's choice is random, is the result of measurement itself. The case that Bob chooses a predetermined value of \mathbf{b} will be considered at the end of this section.

State (3), with register B in a sharp state, is the input state of the conventional representation of the quantum algorithm. For reasons that will become clear, we retard to the end of the unitary part of the quantum algorithm the projection of the quantum state induced by measuring \hat{B} – as well known, such projections can be retarded or advanced along the unitary transformation after or before the measurement (by the way, retarding the projection is here equivalent to deferring the measurement itself). Thus, the input state of the quantum algorithm is state (1) back again.

At this point, Alice applies the Hadamard transform U_A to register A :

$$U_A |\psi\rangle = \frac{1}{4\sqrt{2}} (e^{i\varphi_0} |00\rangle_B + e^{i\varphi_1} |01\rangle_B + e^{i\varphi_2} |10\rangle_B + e^{i\varphi_3} |11\rangle_B) \\ (|00\rangle_A + |01\rangle_A + |10\rangle_A + |11\rangle_A) (|0\rangle_V - |1\rangle_V) \quad (4)$$

Then she performs the reversible computation of $\delta(\mathbf{b}, \mathbf{a})$, represented by the unitary transformation U_f (f like "function evaluation"):

$$U_f U_A |\psi\rangle = \frac{1}{4\sqrt{2}} \begin{bmatrix} e^{i\varphi_0} |00\rangle_B (-|00\rangle_A + |01\rangle_A + |10\rangle_A + |11\rangle_A) + \\ e^{i\varphi_1} |01\rangle_B (|00\rangle_A - |01\rangle_A + |10\rangle_A + |11\rangle_A) + \\ e^{i\varphi_2} |10\rangle_B (|00\rangle_A + |01\rangle_A - |10\rangle_A + |11\rangle_A) + \\ e^{i\varphi_3} |11\rangle_B (|00\rangle_A + |01\rangle_A + |10\rangle_A - |11\rangle_A) \end{bmatrix} (|0\rangle_V - |1\rangle_V), \quad (5)$$

We can see that U_f maximally entangles registers B and A . Four orthogonal states of B , each a value of \mathbf{b} , one by one multiply four orthogonal states of A . This means that the information about the value of \mathbf{b} has propagated to register A .

We should note that the entanglement in state (5) does not yet correspond to correlation between measurement outcomes – measuring \hat{B} and \hat{A} in this state (always in the computational basis) would yield uncorrelated outcomes. To make correlation of entanglement, we need to apply to register A the unitary transformation U'_A (the so called *inversion about the mean*):

$$U'_A U_f U_A |\psi\rangle = \frac{1}{2\sqrt{2}} (e^{i\varphi_0} |00\rangle_B |00\rangle_A + e^{i\varphi_1} |01\rangle_B |01\rangle_A + e^{i\varphi_2} |10\rangle_B |10\rangle_A + e^{i\varphi_3} |11\rangle_B |11\rangle_A) \\ (|0\rangle_V - |1\rangle_V), \quad (6)$$

In state (6), register A contains the solution of the problem – the value of \mathbf{b} chosen by Bob. Alice acquires the solution by measuring \hat{A} . This projects state (6) on:

$$P_A U'_A U_f U_A |\psi\rangle = \frac{1}{\sqrt{2}} |01\rangle_B |01\rangle_A (|0\rangle_V - |1\rangle_V). \quad (7)$$

We note that this projection coincides with the (retarded) projection due to the measurement of \hat{B} in state (1).

Equations (1) and (4) through (7) represent the quantum algorithm *relativized* to the observer Alice in the sense of relational quantum mechanics. By the definition of the problem, the projection due to measuring \hat{B} in state (1) should remain hidden to the observer Alice. It should in fact be retarded until Alice measures \hat{A} in state (6). In view of what will follow, it is important to note that, in the relativized algorithm, the two bit entropy of state (1) represents Alice's initial ignorance of Bob's choice. When Alice measures \hat{A} in state (6), the entropy of the quantum state becomes zero and Alice acquires full knowledge of Bob's choice. Thus, the entropy of the quantum state along the relativized algorithm gauges Alice's ignorance of Bob's choice.

We can see that there is quantum correlation between the outcome of measuring \hat{B} in state (1) and that of measuring \hat{A} in state (6). In fact the two

measurements provide a common, randomly selected, eigenvalue (the value 01 of both \mathbf{b} and \mathbf{a} , i. e. Bob's choice). As we will see, this quantum correlation plays a crucial role in the present explanation of the speed-up.

Until now we have assumed that Bob's choice is a random measurement outcome. An equally crucial point of our argument is noting that quantum correlation remains there also when Bob chooses a predetermined value of \mathbf{b} . Say that the measurement of \hat{B} in state (1) yields $\mathbf{b} = 11$ and Bob wants $\mathbf{b} = 01$. He should apply to register B a permutation of the values of \mathbf{b} , a unitary transformation U_B such that $U_B |11\rangle_B = |01\rangle_B$. Then the algorithm proceeds as before. The correlation is the same as before up to the permutation U_B . The point is that, from the standpoint of quantum correlation, U_B should be considered a "fixed" transformation.

In fact quantum correlation concerns two measurement outcomes in an ensemble of repetitions of the same quantum experiment, here consisting of the measurement of an observable in an initial state, a successive unitary transformation, and the measurement of another observable in the resulting state – Figure 1. Thus, initial state and unitary transformation should remain unaltered throughout the ensemble of repetitions. In particular U_B , being part of the unitary transformation, should be considered the same in all repetitions.

Thus, from the standpoint of quantum correlation, the predetermined value of \mathbf{b} , being the fixed permutation of a random measurement outcome, is a random measurement outcome as well. The fact that Bob chooses U_B to always obtain $\mathbf{b} = 01$, independently of the outcome of measuring \hat{B} in state (1), is independent of quantum correlation.

3.2 Sharing the determination of Bob's choice

We want to share the determination of Bob's choice between Bob's and Alice's measurements.

To start with, we introduce the ingredients required to perform the sharing.

We call $|\psi\rangle_B$ the state of register B at the end of the unitary part of the algorithm, namely in state (6):

$$|\psi\rangle_B = \frac{1}{2} (e^{i\varphi_0} |00\rangle_B + e^{i\varphi_1} |01\rangle_B + e^{i\varphi_2} |10\rangle_B + e^{i\varphi_3} |11\rangle_B). \quad (8)$$

$|\psi\rangle_B$ is the random phase representation of the reduced density operator of register B :

$$\langle\psi\rangle_B \langle\psi|_B\rangle_{\forall\varphi_i} = \frac{1}{4} (|00\rangle_B \langle 00|_B + |01\rangle_B \langle 01|_B + |10\rangle_B \langle 10|_B + |11\rangle_B \langle 11|_B).$$

Incidentally, we note that the state of register B is the same in state (1); in fact it remains unaltered throughout the unitary part of the quantum algorithm which is the identity on it. \mathcal{E}_B , the entropy of $|\psi\rangle_B$, is two bits. The measurement of \hat{B} reduces the entropy of both the state of register B and the overall quantum state by \mathcal{E}_B .

The determination of Bob's choice is represented by P_B , the projection of $|\psi\rangle_B$ on $|01\rangle_B$ due to the measurement of the content of register B . We share the determination of Bob's choice by sharing P_B , what can be done by resorting to the notion of partial measurement of the content of B .

We consider the following partial measurements and the corresponding projections of $|\psi\rangle_B$. The measurement of the content of the left cell of B – of the observable \hat{B}_0 of eigenvalue b_0 . A-priori, the measurement outcome is either $b_0 = 0$ or $b_0 = 1$. However, in present assumptions, the measurement of \hat{B} projects $|\psi\rangle_B$ on $|01\rangle_B$, we are in fact discussing how to share this projection. Thus we should assume that the measurement of \hat{B}_0 yields $b_0 = 0$, namely projects $|\psi\rangle_B$ on $\frac{1}{\sqrt{2}}(e^{i\varphi_0}|00\rangle_B + e^{i\varphi_1}|01\rangle_B)$; we also say "on $\mathbf{b} \in \{01, 00\}$ ". Similarly, the measurement of the content of the right cell of B projects $|\psi\rangle_B$ on $\mathbf{b} \in \{01, 11\}$, that of the exclusive or of the contents of the two cells projects $|\psi\rangle_B$ on $\mathbf{b} \in \{01, 10\}$.

We will see afterwards that P_B should be shared into any two of the three projections of $|\psi\rangle_B$ on: $\mathbf{b} \in \{01, 00\}$, $\mathbf{b} \in \{01, 11\}$, and $\mathbf{b} \in \{01, 10\}$. One share (either one) should be ascribed to the measurement of Bob, the other to that of Alice.

Until now we have introduced the tools to share the determination of Bob's choice between Bob's and Alice's measurements. Now we introduce some conditions that reasonably should be satisfied by the sharing.

First, we get rid of all redundancy between the two measurements. We resort to Occam's razor; in Newton's formulation, it states "*We are to admit no more causes of natural things than such that are both true and sufficient to explain their appearances*" [11]. This requires that, together, the two shares of P_B (the corresponding partial measurements) *tightly* determine the value of \mathbf{b} , namely without determining twice any Boolean function of \mathbf{b} . This is condition (i) of the *sharing rule*.

We apply it to Grover's algorithm. Here, the (more in general) n bits that specify the value of \mathbf{b} are independently selected in a random way. Thus, condition (i) requires that the determination of p of these bits ($0 \leq p \leq n$) is ascribed to the measurement of Bob, that of the other $n - p$ bits to the measurement of Alice.

Condition (i) does not constrain the value of p . This is up to the following condition (ii). Let $\Delta\mathcal{E}_B^{(B)}$ ($\Delta\mathcal{E}_B^{(A)}$) be the reduction of the entropy of the state of register B – and of the overall quantum state – associated with the share of P_B ascribed to Bob's (Alice's) measurement. Here we have $\Delta\mathcal{E}_B^{(B)} = p$, $\Delta\mathcal{E}_B^{(A)} = n - p$. Since Bob's choice is determined by either Bob's or Alice's measurement, for reasons of symmetry we require:

$$\Delta\mathcal{E}_B^{(B)} = \Delta\mathcal{E}_B^{(A)}. \quad (9)$$

Here this becomes $p = n - p = n/2$ – the n bits of \mathcal{E}_B share evenly between the two measurements.

Although self evident, the sharing rule we are dealing with is axiomatic in character and thus open to criticism. However, the present work has an intrinsic

redundancy. The mechanism of the quantum speed-up the sharing rule leads to holds per se, is a common pattern of all the quantum algorithms of the present family. Only the fundamental justification of this pattern relies on the axiomatic step.

We can see that sharing P_B into any two of the above said three projections satisfies conditions (i) and (ii). In fact: (i) any pair of projections, corresponding to the measurement of a pair of observables among \hat{B}_0 , \hat{B}_1 , and \hat{B}_X , tightly selects a value of \mathbf{b} and (ii) any projection reduces the entropy of the state of register B of one bit, so that Eq. (9) is always satisfied. We can also see that there is no other way of satisfying the sharing rule.

Sharing between the measurements of \hat{A} and \hat{B} the determination of Bob's choice is equivalent to saying that Alice's measurement of \hat{A} contributes to this determination. Thus, in Grover's algorithm, Alice's measurement determines half of the bits that specify Bob's choice.

This faces us with the problem that half of Bob's choice can be taken in many ways. A natural way of solving this problem is requiring that the sharing is done in a quantum superposition of the possible ways of taking half of the choice. This is condition (iii) of the sharing rule. We reconcile the quantum algorithm with this condition by assuming that it is a quantum superposition of algorithms (or "histories"), each characterized by a possible way of taking half choice.

3.3 Advanced knowledge

We show that ascribing to Alice's measurement the determination of part of Bob's choice implies that Alice knows in advance, before running the algorithm, that part of the choice.

We work on an example. We ascribe to Alice's measurement the determination $b_0 = 0$ – namely the projection of state (6) on:

$$\frac{1}{2} (e^{i\varphi_0} |00\rangle_B |00\rangle_A + e^{i\varphi_1} |01\rangle_B |01\rangle_A) (|0\rangle_V - |1\rangle_V). \quad (10)$$

To be seen as a contribution to Bob's choice, this projection should be advanced to the time of Bob's choice – see Figure 1. To this end, we apply $U_A^\dagger U_f^\dagger U_A^\dagger$ to the two ends of it, namely to states (6) and (10). This yields the projection of the input state of the quantum algorithm (1) on:

$$\frac{1}{2} (e^{i\varphi_0} |00\rangle_B + e^{i\varphi_1} |01\rangle_B) |00\rangle_A (|0\rangle_V - |1\rangle_V). \quad (11)$$

Thus, the entropy of the state of register B in the input state of the quantum algorithm is halved. Since this entropy represents Alice's initial ignorance of Bob's choice (Section 2.1), this means that Alice, before running the algorithm, knows $n/2$ of the bits that specify Bob's choice, here one bit – in fact $b_0 = 0$.

We are at the level of elementary logical operations where knowing is doing. "Doing" means identifying the missing half of the choice with a single computation of $\delta(\mathbf{b}, \mathbf{a})$. This fits advanced knowledge and explains the speed-up from three to one computation.

3.4 History superposition picture

We show that Grover's algorithm is a quantum superposition of histories; in each of them, given the advanced knowledge of one bit of Bob's choice, Alice classically computes the missing bit.

Let us assume that Bob's choice is $\mathbf{b} = 01$. As we have seen in Section 2.2, Alice's advanced knowledge can be: $\mathbf{b} \in \{01, 00\}$, or $\mathbf{b} \in \{01, 11\}$, or $\mathbf{b} \in \{01, 10\}$.

We start with the first possibility. Given the advanced knowledge of $\mathbf{b} \in \{01, 00\}$, to identify the value of \mathbf{b} Alice should compute $\delta(\mathbf{b}, \mathbf{a})$ (for short " δ ") for either $\mathbf{a} = 01$ or $\mathbf{a} = 00$.

Let us assume it is for $\mathbf{a} = 01$. The outcome of the computation is $\delta = 1$, which of course tells Alice that $\mathbf{b} = 01$. This corresponds to two classical computation histories, one for each possible sharp state of register V : we represent each classical computation history as a sequence of sharp quantum states.

The initial state of history 1 is $e^{i\varphi_1} |01\rangle_B |01\rangle_A |0\rangle_V$. The product $|01\rangle_B |01\rangle_A$ means that the input of the computation of $\delta(\mathbf{b}, \mathbf{a})$ is $\mathbf{b} = 01$, $\mathbf{a} = 01$; $|0\rangle_V$ is one of the two possible sharp states of register V . The state after the computation of δ is $e^{i\varphi_1} |01\rangle_B |01\rangle_A |1\rangle_V$ – the result of the computation is modulo 2 added to the former content of register V . We are using the history amplitudes that reconstruct the quantum algorithm; our present aim is to show that the quantum algorithm is a superposition of histories whose computational part is classical.

By the way, history amplitudes – in equivalent terms the state of registers A and V in (4) – can also be found from scratch by starting with a generic state of the two registers and then setting the amplitudes in such a way that the entanglement between registers B and A (or the entropy of the state of register A) after the first computation of δ is maximized – see Ref. [8].

In history 2, the states before/after the computation of δ are $-e^{i\varphi_1} |01\rangle_B |01\rangle_A |1\rangle_V \rightarrow -e^{i\varphi_1} |01\rangle_B |01\rangle_A |0\rangle_V$.

In the case that Alice computes $\delta(\mathbf{b}, \mathbf{a})$ for $\mathbf{a} = 00$ instead, she obtains $\delta = 0$, which of course tells her again that $\mathbf{b} = 01$. This originates other two histories. History 3: $e^{i\varphi_1} |01\rangle_B |00\rangle_A |0\rangle_V \rightarrow e^{i\varphi_1} |01\rangle_B |00\rangle_A |0\rangle_V$; history 4: $-e^{i\varphi_1} |01\rangle_B |00\rangle_A |1\rangle_V \rightarrow -e^{i\varphi_1} |01\rangle_B |00\rangle_A |1\rangle_V$.

We develop in a similar way the other histories. The computation step of Grover's algorithm, namely the transformation of state (4) into state (5), is the superposition of all these histories.

At this point we perform a non-computational step (the inversion about the mean) by applying the unitary transformation U'_A to register A . This branches each history into four histories; the end states of such branches interfere with one another to give state (6).

Summing up, Grover's algorithm for $n = 2$ can be decomposed into a superposition of histories which start from Alice's advanced knowledge and whose computational part is entirely classical.

Let us now consider the case $n > 2$. As well known, the sequence "function evaluation-inversion about the mean" (the algorithm's *iterate*) should be iterated $\frac{\pi}{4} 2^{n/2}$ times. This maximizes the probability of finding the solution leaving

a probability of error $\leq \frac{1}{2^n}$. This goes along with the present explanation of the speed-up in the order of magnitude. In fact, according to it, one should perform $O(2^{n/2})$ computations of δ – this is the number of classical computations required to find the missing half of Bob’s choice given the advanced knowledge of the other half.

4 Mechanism of the quantum speed-up

The present explanation allows us to pinpoint the exact mechanism of the quantum speed-up. We represent this mechanism by means of a *generalized quantum algorithm* that, given any set of functions and under maximization of a quantity of information, becomes the quantum algorithm that yields with a speed-up a characteristic of the function chosen by Bob.

We start from some set of functions $f_{\mathbf{b}}(\mathbf{a})$, where \mathbf{b} is a bit string ranging over some set of values. In the case of Grover’s algorithm, this is the set of all the $f_{\mathbf{b}}(\mathbf{a}) \equiv \delta(\mathbf{b}, \mathbf{a})$ for $\mathbf{b} \in \{0, 1\}^n$. Register B contains \mathbf{b} , the label of the function chosen by Bob, register A the argument of the function, and register V the result of function evaluation reversibly added to its former content. As we will see, the problem solved by the generalized algorithm should be identified a-posteriori. The algorithm is:

(I) Assume that B is in a maximally mixed state, prepare A and V in such a way that the entanglement between B and A after the first function evaluation is maximum. In equivalent terms the entropy of the state (i. e. of the reduced density operator) of register A should be maximized.

(II) Perform function evaluation. Since the state of register B remains unaltered throughout the unitary part of the algorithm, the creation of entanglement between B and A means that information about the content of B leaks to A . The amount of the information leaked is the entropy of the state of register A . The information is naturally about some characteristic of $f_{\mathbf{b}}(\mathbf{a})$ – the value of \mathbf{b} in Grover’s algorithm or something related to the period of the function in the algorithms of Simon and Shor.

(III) Apply to A a unitary transformation that makes correlation of entanglement. Points (II) and (III) constitute the algorithm’s *iterate*.

(IV) Let \mathcal{N} be the number of function evaluations required to classically determine Bob’s choice given the advanced knowledge of the part of it established by the sharing rule. Perform the iterate $O(\mathcal{N})$ times. This makes the characteristic of $f_{\mathbf{b}}(\mathbf{a})$, itself a function of Bob’s choice, determined in all the quantum algorithms examined in this paper. We incidentally note that, since the present maximization procedure completely defines each iteration of the quantum algorithm, the number of function evaluations required to obtain the solution with a desired probability could also be found analytically.

(V) Acquire the characteristic by measuring \hat{A} .

Thus, the problem solved by the algorithm is finding the characteristic of the function that leaks to register A with function evaluation. Solving this problem is a by-product of reconstructing Bob’s choice. In Ref. [8], by way

of exemplification, we had provided a new quantum speed-up. This speed-up, given the set of functions considered in [8], is automatically generated by the present mechanism.

It should also be noted that the generalized algorithm sharply diverges from Grover's algorithm if we iterate beyond $O(\mathcal{N})$. Having replaced the inversion about the mean by the transformation that makes correlation of entanglement, it is never the case that we unentangle registers B and A . For example, in the case $n = 2$, this latter transformation is exactly the inversion about the mean at the first iteration and becomes the identity if we iterate more than once.

5 Deutsch&Jozsa's algorithm

In Deutsch&Jozsa's [12] algorithm, the set of functions known to both Bob and Alice is all the constant and *balanced* functions (with an even number of zeroes and ones) $f_{\mathbf{b}} : \{0, 1\}^n \rightarrow \{0, 1\}$. Array (12) gives this set for $n = 2$.

\mathbf{a}	$f_{0000}(\mathbf{a})$	$f_{1111}(\mathbf{a})$	$f_{0011}(\mathbf{a})$	$f_{1100}(\mathbf{a})$	$f_{0101}(\mathbf{a})$	$f_{1010}(\mathbf{a})$	$f_{0110}(\mathbf{a})$	$f_{1001}(\mathbf{a})$
00	0	1	0	1	0	1	0	1
01	0	1	0	1	1	0	1	0
10	0	1	1	0	0	1	1	0
11	0	1	1	0	1	0	0	1

(12)

The string $\mathbf{b} \equiv b_0, b_1, \dots, b_{2^n-1}$ is both the suffix and the table of the function – the sequence of function values for increasing values of the argument. Specifying the choice of the function by means of the table of the function simplifies the discussion. Alice is to find whether the function selected by Bob is balanced or constant by computing $f_{\mathbf{b}}(\mathbf{a}) \equiv f(\mathbf{b}, \mathbf{a})$ for appropriate values of \mathbf{a} . In the classical case this requires, in the worst case, a number of computations of $f(\mathbf{b}, \mathbf{a})$ exponential in n ; in the quantum case one computation.

5.1 Time symmetric representation

In the time-symmetric representation, the input and output states of the quantum algorithm are respectively:

$$|\psi\rangle = \frac{1}{4} (e^{i\varphi_0} |0000\rangle_B + e^{i\varphi_1} |1111\rangle_B + e^{i\varphi_2} |0011\rangle_B + e^{i\varphi_3} |1100\rangle_B + \dots) |00\rangle_A (|0\rangle_V - |1\rangle_V), \quad (13)$$

$$U_A U_f U_A |\psi\rangle = \frac{1}{4} [(e^{i\varphi_0} |0000\rangle_B - e^{i\varphi_1} |1111\rangle_B) |00\rangle_A + (e^{i\varphi_2} |0011\rangle_B - e^{i\varphi_3} |1100\rangle_B) |10\rangle_A + \dots] (|0\rangle_V - |1\rangle_V). \quad (14)$$

U_A is the Hadamard transform on register A , U_f is function evaluation, namely the computation of $f(\mathbf{b}, \mathbf{a})$. Measuring \hat{B} in state (13) yields Bob's choice, a value of \mathbf{b} . Measuring \hat{A} in state (14) yields the solution: "constant" if \mathbf{a} is all zeros, "balanced" otherwise.

This time the solution found by Alice is a function of Bob's choice with a lower information content. However – as we will see – this solution is still a by-product of Alice's reconstruction of Bob's choice. This can be explicitly represented by adding an imaginary register \mathcal{V} of the same size of B . The black box, besides reversibly writing in V the result of function evaluation, should reversibly write in \mathcal{V} the corresponding reconstruction of Bob's choice. States (13) and (14) are replaced by

$$\frac{1}{4} (e^{i\varphi_0} |0000\rangle_B + e^{i\varphi_1} |1111\rangle_B + \dots) |0000\rangle_{\mathcal{V}} |00\rangle_A (|0\rangle_V - |1\rangle_V) \quad (15)$$

and

$$\frac{1}{4} [e^{i\varphi_0} (|0000\rangle_B |0000\rangle_{\mathcal{V}} - e^{i\varphi_1} |1111\rangle_B |1111\rangle_{\mathcal{V}}) |00\rangle_A + \dots] (|0\rangle_V - |1\rangle_V). \quad (16)$$

Since all computations are logically reversible, there is certainly a unitary transformation that (with one function evaluation) sends state (15) into (16).

5.2 Sharing the determination of Bob's choice

We should think that Alice, besides \hat{A} , also measures $\hat{\mathcal{V}}$ – the content of register \mathcal{V} . At this point we can share the determination of Bob's choice between the measurement of \hat{B} and that of $\hat{\mathcal{V}}$ exactly as we did for Grover's algorithm. The fact that the measurement of $\hat{\mathcal{V}}$ is an imaginary operation is irrelevant; retro-causality serves in fact to highlight the mechanism of the quantum speed-up hosted in the unitary part of the algorithm.

The state of registers B in the overall state (14) is:

$$|\psi\rangle_B = \frac{1}{2\sqrt{2}} (e^{i\varphi_0} |0000\rangle_B + e^{i\varphi_1} |1111\rangle_B + e^{i\varphi_2} |0011\rangle_B + e^{i\varphi_3} |1100\rangle_B + \dots). \quad (17)$$

We assume that Bob's choice is $\mathbf{b} = 0011$. P_B is the projection of $|\psi\rangle_B$ on $|0011\rangle_B$. \mathcal{E}_B is the (three bit) entropy of $|\psi\rangle_B$. It is zeroed by the measurement of either \hat{B} or $\hat{\mathcal{V}}$.

To share P_B , we consider the "elementary" one-bit observables \hat{B}_i , with $i \in \{0, 1\}^n$. The measurement of each \hat{B}_i yields the value the i -th bit of the bit string \mathbf{b} , namely of the i -th row of the table of the function (as for point **1** of Section 2.2, measurement outcomes are post-selected to match with the bit string). Now that \mathbf{b} is structured we do not need to consider Boolean functions of the \hat{B}_i – we will see that all the histories are generated without such dependent observables (considering also these observables would simply generate repeated histories).

Let us call $P_{B,i}$ the projection of $|\psi\rangle_B$ associated with the measurement of \hat{B}_i . Since each share of P_B is an aggregate of the $P_{B,i}$, it is completely defined by the share of the table on which it projects. Thus, we should take two shares of the table such that the projections on them satisfy the sharing rule.

As shown below, the sharing rule implies that the two shares are two complementary half tables in each of which all the values of the function are the same. We call each share of this kind a *good half table*.

For example in the case $\mathbf{b} = 0011$, these two shares are: $f_{\mathbf{b}}(00) = 0, f_{\mathbf{b}}(01) = 0$ and $f_{\mathbf{b}}(10) = 1, f_{\mathbf{b}}(11) = 1$. The corresponding shares of P_B are the projections of $|\psi\rangle_B$ on $\mathbf{b} \in \{0011, 0000\}$ and respectively $\mathbf{b} \in \{0011, 1111\}$; we have $\Delta\mathcal{E}_B^{(B)} = \Delta\mathcal{E}_B^{(A)} = 2$ bit, what satisfies Eq. (9). One projection (either one) is ascribed to the measurement of \hat{B} , the other to that of \hat{V} . This latter projection becomes Alice knowing in advance the good half table.

We can see that this is the only way of satisfying the sharing rule. If one half table is not good (the values of the function are not all the same), because of the structure of the table, also the other half would not be good. Thus, the two corresponding shares of P_B would both determine the fact that the function is balanced, namely a Boolean function of \mathbf{b} . This would violate the no over-determination condition of the sharing rule. If one or both shares were less than half table, this would either not satisfy Eq. (9) or not determine the value of \mathbf{b} , as readily checked.

5.3 Advanced knowledge

Besides Alice's contribution to Bob's choice, a good half table represents Alice's advanced knowledge of Bob's choice. In fact, since the state of register B remains unaltered throughout the unitary part of the quantum algorithm, also its projection on the half table remains unaltered. At the end of the unitary part of the algorithm, this projection represents Alice's contribution to Bob's choice. Advanced at the beginning, it changes Alice's complete ignorance of Bob's choice into knowledge of the half table.

It is immediate to check that the quantum algorithm requires the number of function evaluations of a classical algorithm that has to identify Bob's choice starting from the advanced knowledge of a good half table. In fact, the value of \mathbf{b} (and the solution of the problem) can be identified by computing $f_{\mathbf{b}}(\mathbf{a})$ for only one value of \mathbf{a} (anyone) outside the half table. Thus, both the quantum algorithm and the advanced knowledge classical algorithm require just one function evaluation.

5.4 History superposition picture

It is convenient to group the histories with the same value of \mathbf{b} . Starting with $\mathbf{b} = 0011$, we assume that Alice's advanced knowledge is, e. g., $\mathbf{b} \in \{0011, 0000\}$. In order to determine the value of \mathbf{b} and thus the character of the function, Alice should perform function evaluation for either $\mathbf{a} = 10$ or $\mathbf{a} = 11$. We assume it is for $\mathbf{a} = 10$. Since we are under the assumption $\mathbf{b} = 0011$, the result of the computation is 1. This originates two classical computation histories, each consisting of a state before and one after function evaluation. History 1: $e^{i\varphi_2} |0011\rangle_B |10\rangle_A |0\rangle_V \rightarrow e^{i\varphi_2} |0011\rangle_B |10\rangle_A |1\rangle_V$. His-

tory 2: $-e^{i\varphi^2} |0011\rangle_B |10\rangle_A |1\rangle_V \rightarrow -e^{i\varphi^2} |0011\rangle_B |10\rangle_A |0\rangle_V$. If she performs function evaluation for $\mathbf{a} = 11$ instead, this originates other two histories, etc.

As readily checked, the superposition of all these histories is the function evaluation stage of the quantum algorithm. Then, Alice applies the Hadamard transform to register A . Each history branches into four histories. The end states of such branches interfere with one another to yield state (14).

We can see that the present quantum algorithm conforms to the generalized algorithm. In each history, Alice performs the computations required to determine Bob's choice (which means just one function evaluation). The solution of Deutsch&Jozsa's problem is a by-product of this determination.

It is easy to see that the present analysis holds unaltered for $n > 2$.

6 Simon's and the hidden subgroup algorithms

In Simon's [13] algorithm, the set of functions is all the $f_{\mathbf{b}} : \{0, 1\}^n \rightarrow \{0, 1\}^{n-1}$ such that $f_{\mathbf{b}}(\mathbf{a}) = f_{\mathbf{b}}(\mathbf{c})$ if and only if $\mathbf{a} = \mathbf{c}$ or $\mathbf{a} = \mathbf{c} \oplus \mathbf{h}^{(\mathbf{b})}$; \oplus denotes bitwise modulo 2 addition; the bit string $\mathbf{h}^{(\mathbf{b})}$, depending on \mathbf{b} and belonging to $\{0, 1\}^n$ excluded the all zeroes string, is a sort of period of the function. Array (18) gives the set of functions for $n = 2$. The bit string \mathbf{b} is both the suffix and the table of the function. Since $\mathbf{h}^{(\mathbf{b})} \oplus \mathbf{h}^{(\mathbf{b})} = \mathbf{0}$ (the all zeros string), each value of the function appears exactly twice in the table, thus 50% of the rows plus one always identify $\mathbf{h}^{(\mathbf{b})}$.

	$\mathbf{h}^{(0011)} = 01$	$\mathbf{h}^{(1100)} = 01$	$\mathbf{h}^{(0101)} = 10$	$\mathbf{h}^{(1010)} = 10$	$\mathbf{h}^{(0110)} = 11$	$\mathbf{h}^{(1001)} = 11$
\mathbf{a}	$f_{0011}(\mathbf{a})$	$f_{1100}(\mathbf{a})$	$f_{0101}(\mathbf{a})$	$f_{1010}(\mathbf{a})$	$f_{0110}(\mathbf{a})$	$f_{1001}(\mathbf{a})$
00	0	1	0	1	0	1
01	0	1	1	0	1	0
10	1	0	0	1	1	0
11	1	0	1	0	0	1

(18)

Bob selects a value of \mathbf{b} . Alice's problem is finding the value of $\mathbf{h}^{(\mathbf{b})}$, "hidden" in $f_{\mathbf{b}}(\mathbf{a})$, by computing $f_{\mathbf{b}}(\mathbf{a}) = f(\mathbf{b}, \mathbf{a})$ for different values of \mathbf{a} . In present knowledge, a classical algorithm requires a number of computations of $f(\mathbf{b}, \mathbf{a})$ exponential in n . The quantum algorithm solves the hard part of this problem, namely finding a string $\mathbf{s}_j^{(\mathbf{b})}$ orthogonal to $\mathbf{h}^{(\mathbf{b})}$, with one computation of $f(\mathbf{b}, \mathbf{a})$; "orthogonal" means that the modulo 2 addition of the bits of the bitwise product of the two strings is zero. There are 2^{n-1} such strings. Running the quantum algorithm yields one of these strings at random (see further below). The quantum algorithm is iterated until finding $n - 1$ different strings. This allows us to find $\mathbf{h}^{(\mathbf{b})}$ by solving a system of modulo 2 linear equations.

6.1 Time symmetric representation

In the time-symmetric representation, the input and output states of the quantum algorithm are respectively:

$$|\psi\rangle = \frac{1}{2\sqrt{6}} (e^{i\varphi_0} |0011\rangle_B + e^{i\varphi_1} |1100\rangle_B + e^{i\varphi_2} |0101\rangle_B + e^{i\varphi_3} |1010\rangle_B + \dots) |00\rangle_A |0\rangle_V. \quad (19)$$

$$U_A U_f U_A |\psi\rangle = \frac{1}{2\sqrt{6}} \left\{ \begin{array}{l} (e^{i\varphi_0} |0011\rangle_B + e^{i\varphi_1} |1100\rangle_B) [(|00\rangle_A + |10\rangle_A) |0\rangle_V + (|00\rangle_A - |10\rangle_A) |1\rangle_V] + \\ (e^{i\varphi_2} |0101\rangle_B + e^{i\varphi_3} |1010\rangle_B) [(|00\rangle_A + |01\rangle_A) |0\rangle_V + (|00\rangle_A - |01\rangle_A) |1\rangle_V] + \dots \end{array} \right\} \quad (20)$$

In state (19), register V is prepared in the all zeros string (just one zero for $n = 2$). State (20) is reached with a single computation of $f(\mathbf{b}, \mathbf{a})$. In state (20), for each value of \mathbf{b} , register A (no matter the content of V) hosts even weighted superpositions of the 2^{n-1} strings $\mathbf{s}_j^{(\mathbf{b})}$ orthogonal to $\mathbf{h}^{(\mathbf{b})}$. By measuring \hat{A} in this state, Alice obtains at random one of the $\mathbf{s}_j^{(\mathbf{b})}$. Then she iterates the "right part" of the algorithm (preparation of registers A and V , computation of $f(\mathbf{b}, \mathbf{a})$, and measurement of \hat{A}) until obtaining $n - 1$ different $\mathbf{s}_j^{(\mathbf{b})}$.

By the way, Alice should throw away all the measurement outcomes where the value of \mathbf{a} is 00: such outcomes are completely uncorrelated with the value of \mathbf{b} – see Eq. (20) – are thus cases where the quantum algorithm fails. We note that the probability of getting the measurement outcome $\mathbf{a} = 00$ is $1/2^{n-1}$. From now on we consider this "post-selected" quantum algorithm.

6.2 Sharing the determination of Bob's choice

Also in the present case the solution of the problem is a by-product of reconstructing Bob's choice. The analysis is completely similar to that of Section 5.2.

We assume that Bob's choice is $\mathbf{b} = 0011$. P_B is the projection of $|\psi\rangle_B$ on $|0011\rangle_B$. The shares of P_B are still defined by the shares of the table 0011 on which they project. This time a good half table should not contain a same value of the function twice, what would over-determine $\mathbf{h}^{(\mathbf{b})}$, namely a Boolean function of \mathbf{b} (because of the structure of the table, also the other half would contain a same value twice). There are two ways of sharing the table 0011. One is $f_{\mathbf{b}}(00) = 0, f_{\mathbf{b}}(10) = 1$ and $f_{\mathbf{b}}(01) = 0, f_{\mathbf{b}}(11) = 1$; the corresponding shares of P_B are the projections of $|\psi\rangle_B$ on $\mathbf{b} \in \{0011, 0110\}$ and $\mathbf{b} \in \{0011, 1001\}$, see array (18). The other is $f_{\mathbf{b}}(00) = 0, f_{\mathbf{b}}(11) = 1$ and $f_{\mathbf{b}}(01) = 0, f_{\mathbf{b}}(10) = 1$, etc.

By the way, we should note that sharing each table into two halves is accidental. In the quantum part of Shor's [14] factorization algorithm (finding the period of a periodic function), taking two parts of the table that do not contain a same value of the function twice implies that each part is less than half table when the domain of the function spans more than two periods.

6.3 Advanced knowledge

Back to Simon's algorithm, we can see that it requires the number of function evaluations of a classical algorithm that knows in advance a good half table. In fact, since no value of the function appears twice in the half table, the value of \mathbf{b} is always determined by computing $f(\mathbf{b}, \mathbf{a})$ for only one value of \mathbf{a} (anyone) outside the half table – see array (18). Thus, both the quantum algorithm and the advanced knowledge classical algorithm require just one function evaluation.

6.4 History superposition picture

The history superposition picture can be developed as in Section 4.4: given the advanced knowledge of, say, $\mathbf{b} \in \{0011, 0110\}$, in order to determine the value of \mathbf{b} , Alice should perform function evaluation for either $\mathbf{a} = 01$ or $\mathbf{a} = 11$, etc.

The present analysis holds unaltered for $n > 2$. It also applies to the generalized Simon's problem and to the Abelian hidden subgroup problem. In fact the corresponding algorithms are essentially the same as the algorithm that solves Simon's problem. In the hidden subgroup problem, the set of functions $f_{\mathbf{b}} : G \rightarrow W$ map a group G to some finite set W with the property that there exists some subgroup $S \leq G$ such that for any $\mathbf{a}, \mathbf{c} \in G$, $f_{\mathbf{b}}(\mathbf{a}) = f_{\mathbf{b}}(\mathbf{c})$ if and only if $\mathbf{a} + S = \mathbf{c} + S$. The problem is to find the hidden subgroup S by computing $f_{\mathbf{b}}(\mathbf{a})$ for various values of \mathbf{a} .

Now, a large variety of problems solvable with a quantum speed-up can be reformulated in terms of the hidden subgroup problem [15]. Among these we find: the seminal Deutsch's problem, finding orders, finding the period of a function (thus the problem solved by the quantum part of Shor's factorization algorithm), discrete logarithms in any group, hidden linear functions, self shift equivalent polynomials, Abelian stabilizer problem, graph automorphism problem.

7 Conclusions

The time-symmetric representation of quantum algorithms shows that the solution of the problem on the part of Alice is a by-product of her reconstruction of Bob's choice. Because of maximal quantum correlation between Bob's choice and its reconstruction, all is as if Alice, by reading the reconstruction at the end of the algorithm, determined half of Bob's choice. This quantum retroaction of the output on the input explains the quantum speed-up, the fact that quantum algorithms create a stronger than classical input-output correlation. This finding:

(I) Allows us to pinpoint the exact mechanism of the quantum speed-up. Given a set of functions and under maximization of a quantity of information, this mechanism generates the quantum algorithm that finds with a speed-up a characteristic of the function chosen by Bob. The number of function evaluations is that required to classically determine Bob's choice given the advanced knowledge of half of it. This should be a break through, until now there was no

known mechanism for producing speed-ups. In hindsight, we can see that the new speed-up provided in Reference [8] is generated by this mechanism.

(II) Highlights an essential difference between quantum and classical causality, showing that causal quantum processes can host a loop of classical causality. The causal quantum process is for example the transformation of $|\Psi\rangle = \frac{1}{\sqrt{2}}|01\rangle_B|00\rangle_A(|0\rangle_V - |1\rangle_V)$ into $U'_A U_f U_A |\Psi\rangle = \frac{1}{\sqrt{2}}|01\rangle_B|01\rangle_A(|0\rangle_V - |1\rangle_V)$ – equations (3) and (7) of Grover’s algorithm. This transformation is a quantum superposition of classical computation histories. In each of them, Alice knows in advance half of Bob’s choice and solves the problem more quickly by computing only the missing half. This partial knowledge of the result of a computation before performing it (a causality loop) would be impossible if histories were isolated with respect to one another. However, this impossibility argument cannot be applied to the present case. In the superposition of all histories, the half choice known in advance in one history becomes the missing half in another one, where it is computed. Thus, all the possible halves of Bob’s choice are computed, in quantum superposition. Moreover, histories are not isolated with respect to one another, as quantum interference provides cross talk between them. The fact that the interplay between quantum superposition and interference can originate loops of classical causality is what explains the quantum speed-up at a fundamental physical level. In particular, it explains why quantum algorithms can violate a limit demonstrably applying to any classical process.

One can envisage two directions of further research. One is the search for new quantum speed-ups. Under the above said mechanism, the question becomes whether there are sets of functions that provide practically interesting speed-ups beyond those already discovered. The second one is the possibility of extending the present explanation of the speed-up to other classes of quantum algorithms. As it is, the explanation applies to problem-solving, it requires seeing a problem in the input of the quantum process and the solution of the problem in the output. This can be difficult in the quantum algorithms based on random walks or in mixed state quantum computing. In some cases, the quantum speed-up is only visible in the fact that the classical simulation of the input-output process requires a higher amount of computational resources. In order to apply the present explanation to these latter algorithms, we should decouple it from problem-solving. We note that the kernel of the explanation already is decoupled, is the fact that the quantum retroaction of the output on the input allows the creation of an input-output correlation stronger than what is classically possible. One can conjecture that this quantum retroaction mechanism stands at the basis of any quantum speed-up.

Acknowledgments

Thanks are due to Pablo Arrighi, Vint Cerf, David Deutsch, Artur Ekert, David Finkelstein, Hartmut Neven, and Daniel Sheehan, for useful comments/discussions.

References

- [1] D. Deutsch, Proc. Roy. Soc. London A **400**, 97 (1985).
- [2] L. K. Grover, in *Proc. of the 28th Annual ACM Symposium on the Theory of Computing, May 22-24* (ACM Press New York, 1996) p. 219.
- [3] Y. Aharonov, P. G. Bergmann and J. Lebowitz, Phys. Rev. B **134**, 1410 (1964).
- [4] L. Vaidman, *Compendium of Quantum Physics: Concepts, Experiments, History and Philosophy* Greenberger, D., Hentschel, K., and Weinert, F., eds. (Springer-Verlag, Berlin Heidelberg 2009).
- [5] S. Dolev and A. C. Elitzur, *Non-sequential behavior of the wave function*, arXiv:quant-ph/0102109v1 (2001).
- [6] C. Rovelli, Int. J. Theor. Phys. **35**, 8, 1637 (1996).
- [7] G. Castagnoli, in *Proc. of the 92nd Annual Meeting of the AAAS Pacific Division, Quantum Retrocausation: Theory and Experiment, San Diego, June 2011* (2011).
- [8] G. Castagnoli, Phys. Rev. A, **82**, Issue 5, 052334 (2010).
- [9] F. Morikoshi, Int. J. Theor. Phys, **50**, 1858 (2011).
- [10] D. Bohm and D. Pines, Phys. Rev. **92**, 609 (1953).
- [11] S. Hawking, *On the Shoulders of Giants*. Running Press, Philadelphia-London, p. 731 (2003).
- [12] D. Deutsch and R. Jozsa, Proc. R. Soc. London A, **439**, 553 (1992).
- [13] D. Simon, in *Proceedings of the 35th Annual IEEE Symposium on the Foundations of Computer Science* (IEEE Computer Society Press, 1994), p. 116.
- [14] P. W. Shor, in *Proc. of the 35th Annual IEEE Symposium on the Foundations of Computer Science* (IEEE Computer Society Press, 1994) p. 124.
- [15] P. Kaye, R. Laflamme, and M. Mosca, *An introduction to Quantum Computing* (Oxford University Press 2007) p. 146.